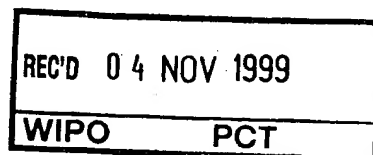


BUNDESREPUBLIK DEUTSCHLAND

09/807638

**PRIORITY  
DOCUMENT**SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)**Bescheinigung**

Die VOLKSWAGEN Aktiengesellschaft in Wolfsburg/Deutschland hat eine  
Patentanmeldung unter der Bezeichnung

"Verfahren und Vorrichtung zur automatischen Steuerung min-  
destens eines Gerätes per Sprachdialog"

am 16. Oktober 1998 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprüng-  
lichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig das Symbol  
G 10 L 7/08 der Internationalen Patentklassifikation erhalten.

München, den 12. Oktober 1999

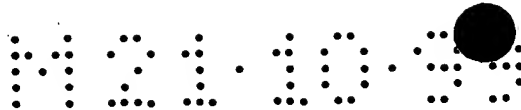
Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

Brand

Aktenzeichen: 198 47 879.8

**VOLKSWAGEN**

K7341/1770-hk-de

## **Verfahren und Vorrichtung zur automatischen Steuerung mindestens eines Gerätes per Sprachdialog**

Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur automatischen Steuerung mindestens eines Gerätes per Sprachdialog.

Verfahren bzw. Vorrichtungen dieser Art werden allgemein in sogenannten Sprachdialog- bzw. Sprachbediensystemen z.B. für Fahrzeuge, Computer, Roboter, Maschinen, Anlagen usw. eingesetzt.

Ein Sprachdialogsystem (SDS) läßt sich i.a. im wesentlichen auf folgende Komponenten reduzieren:

- Spracherkennungssystem, welches ein eingesprochenes Kommando („Sprachkommando“) mit anderen erlaubten Sprachkommandos vergleicht und eine Entscheidung trifft, welches Kommando aller Wahrscheinlichkeit nach eingesprochen wurde,
- Sprachausgabe, welche die zur Benutzerführung erforderlichen Sprachbefehle und Signalisierungstöne ausgibt und ggf. das Erkennungsergebnis rückmeldet,
- Dialog- und Ablaufsteuerung, um dem Benutzer zu verdeutlichen, welche Art von Eingabe erwartet wird, bzw. um zu prüfen, ob die erfolgte Eingabe konsistent ist mit der Aufforderung und mit dem momentanen Status der Applikation, und um die resultierende Aktion bei der Applikation (z.B. dem zu steuernden Gerät) anzustoßen,
- Kontrollinterface als Schnittstelle zur Applikation: Dahinter verbergen sich Hard- und Softwaremodule, um verschiedene Aktuatoren bzw. Rechner anzusteuern, die die Applikation beinhalten,
- Applikation, die per Sprache angesteuert wird: Das kann z.B. ein Bestell- oder Auskunftssystem, ein CAE-Arbeitsplatz oder ein behindertengerechter Rollstuhl sein.

Ein solches Sprach-Dialog-System ist beispielsweise aus der DE 195 33 541 C1 bekannt. Zur Verbesserung der Akzeptanz eines solchen Mensch-Maschinen-Dialogs wird dort beispielsweise vorgeschlagen, für die Kommandos synonyme Worte bzw. unterschiedliche Aussprachevarianten zu verwenden oder aber die Wörter in den Kommandos umzustellen, beispielsweise „Größerer Radius bei linkem Kreis“ oder alternativ hierzu „Bei linkem Kreis

größerer Radius“. Des weiteren wird vorgeschlagen, durch einen erweiterten Speicher ein multilinguales sprecherunabhängiges Dialogsystem aufzubauen, wobei wahlweise zwischen den Dialogsystemen verschiedener Sprachen umgeschaltet werden kann. Weiter wird vorgeschlagen, sogenannte Ellipsen einzubinden, d.h. Verzicht auf Wiederholung ganzer Kommandosätze und statt dessen Kommandos wie „höher“, „stärker“ oder „weiter“, wobei dann das Sprachdialog-System diese den vorangegangenen Kommandos zuordnet. Des weiteren kann das Sprach-Dialog-System bei unsicherem Erkennen Rückfragen wie beispielsweise „Wie bitte?“, „Bitte wiederholen“ oder „Und weiter?“ stellen oder bestimmte Anregungen wie „Bitte lauter“ ausgeben. All diese Maßnahmen dienen dazu, eine monotone Kommunikation zu vermeiden und den Dialog stärker einer Mensch-Mensch-Kommunikation anzunähern. Zur Verbesserung der Kommunikation wird das Sprachsystem mit einem optischen Anzeigemedium gekoppelt, auf dem erkannte Befehle zu Kontrollzwecken angegeben werden. Weiter ermöglicht das optische Anzeigemedium die Darstellung der vom Zielgerät als Reaktion auf das Sprachkommando eingestellten Funktionen und/oder verschiedener Funktionen/Alternativen, die per Sprachkommando anschließend eingestellt bzw. ausgewählt werden können. Nachteilig an der bekannten Vorrichtung und des damit durchgeführten Verfahrens ist, daß trotz der dargestellten Verbesserungen die Sprachausgabe aufgrund ihrer Monotonie den Nutzer ermüdet, so daß dessen Reaktionszeit bei handlungskritischen Aktionen zu groß ist. Ein weiteres Problem ist, daß bei Erkennungsschwierigkeiten die bekannten Sprach-Dialog-Systeme eine Endlosschleife durchlaufen und dem Nutzer immer wieder die gleiche Aufforderung ausgeben, so daß der work-flow unterbrochen ist.

Der Erfindung liegt daher das technische Problem zugrunde, ein Verfahren und eine Vorrichtung zur automatischen Steuerung eines Gerätes per Sprachdialog zu schaffen, bei denen bei handlungskritischen Anweisungen eine verbesserte Aufmerksamkeit des Nutzers erreicht wird sowie der work-flow sichergestellt ist.

Die Lösung des technischen Problems ergibt sich durch die Merkmale der Patentansprüche 1, 2 und 8. Weitere vorteilhafte Ausgestaltungen der Erfindung ergeben sich aus den Unteransprüchen.

Durch die Verwendung unterschiedlicher Intonationen, wobei bei handlungskritischen Anweisungen die Kommandos eine Befehlsintonation aufweisen, wird bereits unmittelbar bei der Sprachausgabe die Aufmerksamkeit des Nutzers geweckt, so daß die Reaktionszeit zur Durchführung der geforderten Anweisung erheblich verkürzt wird.



Zur weiteren Erhöhung der Aufmerksamkeit und Unterscheidung von handlungskritischen Anweisungen kann die Lautstärke der Sprachausgabe bei handlungskritischen Anweisungen erhöht werden und/oder diese besonders hart bzw. abrupt eingeblendet werden.

Weiter kann das Sprach-Dialog-System multilingual ausgebildet sein, so daß beispielsweise zwischen einer Männer- und einer Frauenstimme gewählt werden kann, wobei dann bei handlungskritischen Anweisungen die eine und bei handlungsunkritischen Anweisungen die andere Stimme vom System ausgewählt wird.

Zur Sicherstellung des work-flows wird das Sprach-Dialog-System nur durch Betätigung eines „Push to talk“ PTT-Schalters aktiviert, wobei bei Ausbleiben einer gültigen Interaktion die Dialog-Kommunikations-Ebene gewechselt wird. Zur Erhöhung der Erkennungssicherheit bzw. zur Verbesserung der Nutzerführung können einzelne Kommandos in verschiedenen alternativen Ausgabeformen abgespeichert sein, die dann bei einer ungültigen Interaktion sukzessive ausgegeben werden, wobei die Dialog-Kommunikations-Ebene erst gewechselt wird, wenn auf alle Kommandoformen keine gültige Interaktion erfolgt ist. Zur Vermeidung von Monotonie kann die Reihenfolge der Ausgabe durch einen Zufallsgenerator permutiert werden.

Die Erfindung wird nachfolgend anhand eines bevorzugten Ausführungsbeispiels näher erläutert. Die einzige Figur zeigt ein schematisches Flußdiagramm des Verfahrens zur automatischen Steuerung mindestens eines Gerätes per Sprachdialog.

Durch die Betätigung des PTT-Schalters wird das Sprach-Dialog-System aktiviert. Zum besseren Verständnis wird die Sprachausgabe des Sprachdialog-Systems in Kommandos KOM und Aufforderungen Auff unterteilt, die in Realität identisch sein können. Nachfolgend wird unter Kommandos KOM eine direkte Anweisung zum Handeln verstanden, wie beispielsweise „BREMSSEN“, „LICHTEINSCHALTEN“, wohingegen Aufforderungen eine Interaktion in Form einer Eingabe anfordern, wie beispielsweise „Bitte gewünschte Temperatur in Grad Celsius angeben“.

Erzeugt nun das Sprach-Dialog-System ein Kommando KOM, so wird dieses Kommando KOM danach unterteilt, ob es sich um eine handlungskritische Anweisung oder eine handlungsunkritische Anweisung handelt. Vereinfacht sind handlungskritische Anweisungen, Kommandos KOM, die ein schnelles Durchführen der Handlung erfordern. Ein Beispiel hierfür ist das Kommando KOM „Bremsen“, wenn ein ADR-System oder eine Prescrash-Sensorik ein Kollisionsobjekt erfaßt haben. Handlungsunkritische Anweisungen sind

beispielsweise Kommandos KOM eines Navigationssystems. Dabei werden handlungskritische Anweisungen mit einer Befehlsintonations-Sprache S1, mit großer Lautstärke L1 hart und abrupt in einer Zeit t1 eingeblendet, um beim Nutzer einen hohen Aufmerksamkeitsgrad zu erzeugen. Handlungsunkritische Anweisungen werden hingegen weich bei geringerer Lautstärke L2 und normaler Intonation S2 eingeblendet.

Aufforderungen Auff sind in der Regel zeitunkritisch, so daß hier eine gute Benutzerführung im Vordergrund steht. Dazu kann eine Aufforderung Auff in n verschiedenen Alternativen im Sprachspeicher abgelegt sein. Die Alternativen können beispielsweise unterschiedliche Betonungen, Aussprachen, Wortumstellungen oder synonyme Begriffe sein. Nach akustischer Ausgabe der ersten Alternative wartet das Sprach-Dialog-System eine vorgegebene Zeit auf eine Interaktion. Erfolgt innerhalb dieser Zeit keine oder eine ungültige Interaktion, so wiederholt das Sprach-Dialog-System die Aufforderung mit der zweiten Alternative, gegebenenfalls bis zur n-ten Alternative. Erfolgt eine gültige Interaktion, so wird diese ausgeführt und gegebenenfalls eine neue Aufforderung Auff ausgegeben. Erfolgt hingegen auf die n-te Alternative der Aufforderung Auff keine gültige Interaktion, so wird auf eine andere Dialog-Kommunikations-Ebene DKE gewechselt, um den work-flow sicherzustellen. Die neue Dialog-Kommunikations-Ebene DKE ist dann beispielsweise eine Auswahlliste, die auf dem Bordmonitor dargestellt wird, woraus sich der Nutzer ein entsprechendes Menü auswählen kann.

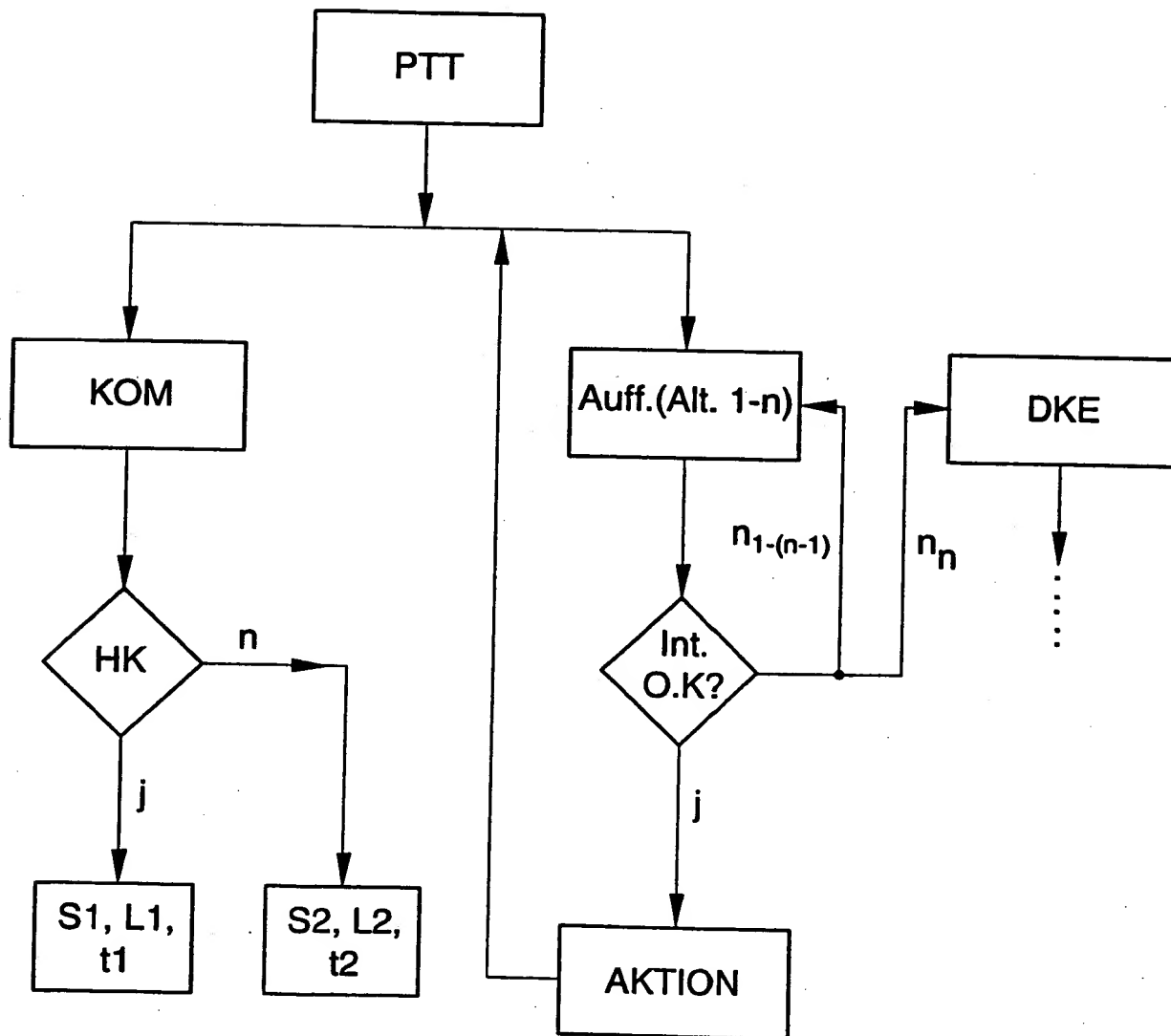


FIG. 1

K7341/1770-hk-de

## **ZUSAMMENFASSUNG**

### **Verfahren und Vorrichtung zur automatischen Steuerung mindestens eines Gerätes per Sprachdialog**

Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur automatischen Steuerung mindestens eines Gerätes per Sprachdialog, umfassend einen Sprachspeicher, in dem für die Sprachausgabe benötigte Kommandos (KOM) in unterschiedlichen Intonationen abgelegt sind, wobei mindestens die Kommandos (KOM) für handlungskritische Anweisungen eine Befehlsintonation aufweisen.

(Fig. 1)

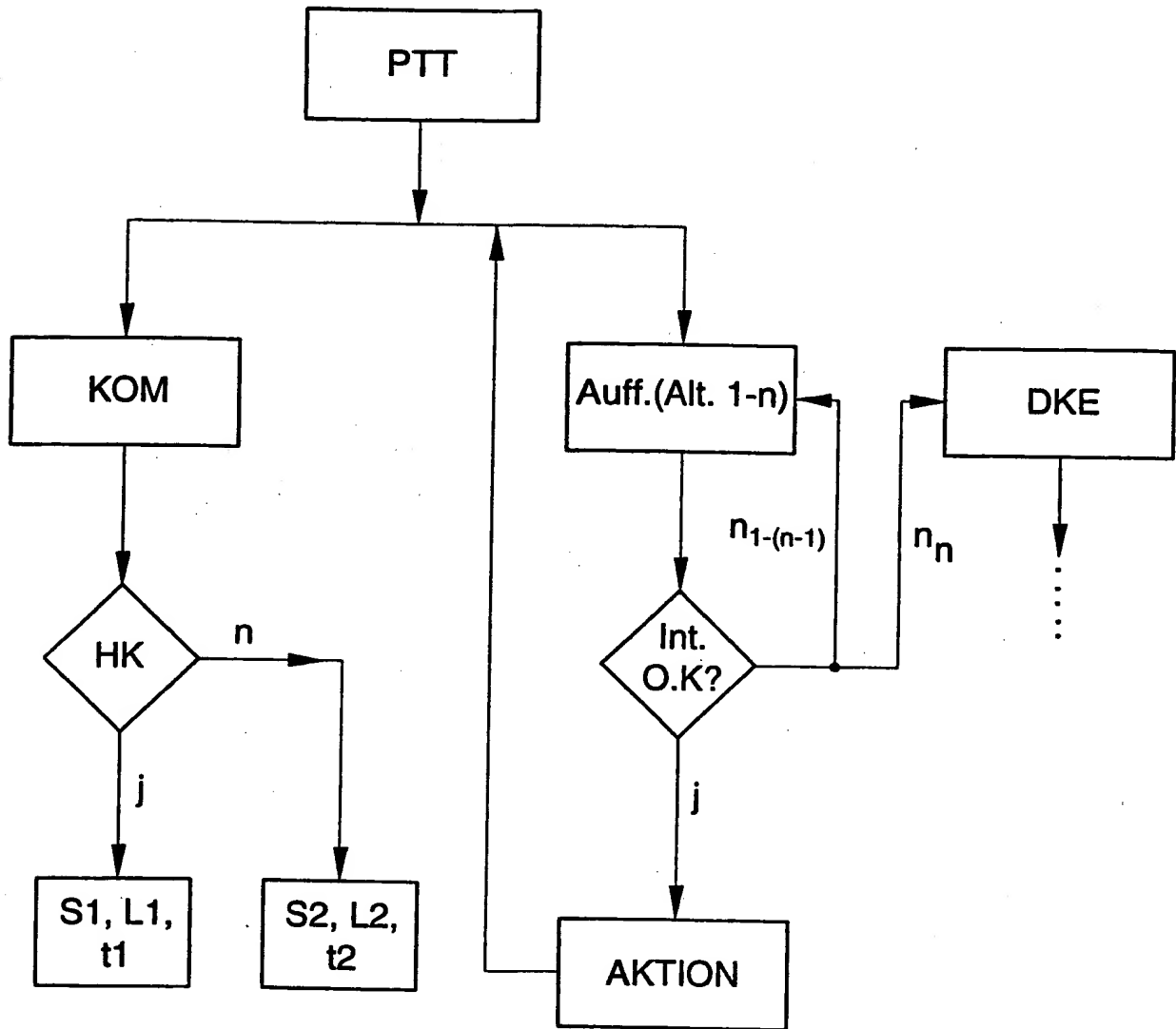


FIG. 1